



Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 163

Agents' Rationality and the CHF/USD Exchange Rate
Part I

Hermann Garbers

October 2003

UNIVERSITY OF ZURICH
INSTITUTE FOR EMPIRICAL RESEARCH IN ECONOMICS
BLÜMLISALPSTRASSE 10, CH-8006 ZURICH

AGENTS' RATIONALITY AND THE CHF/USD EXCHANGE RATE

Part I

Hermann Garbers¹

Version: October 6, 2003

¹ I would like to thank G. Hansen (Kiel), M. Kurz (Stanford), Ch. Bucheli, M. Habib, P. Harpes and C. Ortelli (Zurich) for their helpful and critical discussions. Earlier versions of the paper have been presented in seminars at the University of Zurich, the ETH Zurich, the "Finanzdienstleistungs-Symposium 2002" in Vaduz and the third Workshop in Macroeconometrics at the Halle Institute for Economic Research. Financial support by the *Richard Büchner-Stiftung* is gratefully acknowledged.

Summary

The analysis of monthly exchange rates is carried out using a model of *B.T. McCallum* [12], which is based on the concept of Rational Expectations. Applying the model to the *CHF/USD* exchange rate, s_t , starting a *misspecification analysis*, the *RE* component appears to be a weak point of the model.

The theory of rational beliefs of *M. Kurz* generalizes the *RE* concept introducing special consideration of Data Generating Processes (*DGP*). We find, however, some evidence speaking against the applicability of the rational belief approach (with respect to s_t). It appears that the rationality of economic agents depends on *complex cognitive processes* not discussed by *M. Kurz*, but taken into account in a "story" by *P. De Grauwe* [1]. This story will be supplemented in Part II of the paper in order to proceed with the misspecification analysis of *B.T. McCallum*'s model.

Keywords:

General Economics and Methodology

JEL Classification:

A1, B4, C91, D8, F31

Contents

Introduction	1
1 The McCallum Model	2
2 History of the Model	3
3 Rejecting McCallum's Model	7
4 Rational Beliefs, Stable Processes and Endogeneous Uncertainty	8
5 Switching Regimes in s_t	13
6 A few conclusions	16
References	23

Introduction

It is one of the tasks of an econometrician to determine, or at least approximate, the *DGP* of economic variables. By modeling the variables for instance within the framework of *ARIMA* and *GARCH* processes some progress was made since, say, 1980. But this progress is slightly limited to single variables excluding dollar-based exchange rates. It is, therefore, a challenge to analyse the *DGP* of the *CHF/USD* exchange rate and it makes sense to start the work in a multivariate setting. The question then is which variables should be included into the setting. It seems that one has to look for some kind of economic theory to get an answer. We made a decision in favour of a theory which has been formalized in the following model by *B.T. McCallum*.

But notice – in order to appreciate the problems of modeling exchange rates – a quotation from *R.J. Shiller* [14] concerning the behavior of agents in speculative markets:

”Psychologists have shown that people’s decisions in ambiguous situations are influenced by whatever available anchor is at hand. (...) There are *quantitative* and *moral* anchors. With quantitative anchors, people weigh numbers against prices (e.g. when they decide whether assets are priced right). (...) Underlying the notion of moral anchors is the psychological principle that much of human thinking that results in action is not quantitative but instead takes the form of *storytelling* and *justification* (...). With moral anchoring (...) the market is not prevented from going up to arbitrarily high levels because people have an idea what its intrinsically ”right” level is (...). Rather (...) the discrepancy between the wealth many people would then have in the market and their current living standards would, when compared with their reasons for holding stocks, encourage them to sell (...).

Such reasoning is not well described by the usual kind of economic theory, but there is a large amount of evidence in support (...).”

1 The McCallum Model

Following the introduction we present next the model chosen as our starting point:

$$(I)$$

$$\begin{aligned} s_t &= E_t[s_{t+1}] - z_t + \zeta_t \\ \zeta_t &= \rho \zeta_{t-1} + \varepsilon_t^{(1)} & ; \quad 0 < \rho < 1 \\ z_t &= \lambda(s_t - s_{t-1}) + \gamma z_{t-1} + \varepsilon_t^{(2)} & ; \quad \lambda > 0, \quad 0 < \gamma < 1 \end{aligned}$$

s_t represents here the logarithm of the *CHF/USD* exchange rate, z_t represents an interest rate difference between Switzerland and the USA, ζ_t is a latent variable, which describes, amongst other things, a risk premium. $E_t[s_{t+1}]$ represents a conditional expected value. $\varepsilon_t^{(1)}$, $\varepsilon_t^{(2)}$ are independent "white noise" processes.

The first equation contains a version of the uncovered interest parity with an autoregressive error term of first order, partly and implicitly representing an unobserved time varying risk premium, while the last equation reflects a smoothing of interest rate differentials and a feedback of exchange rates to interest rate differentials. This second relationship is partly due to the fact that the monetary authorities in both countries take short term interest rates as policy instruments to adjust to undesired exchange rate movements.

Taking the conditional expectation of the transformed third equation so that

$$s_t - s_{t-1} = \frac{1}{\lambda} z_t - \frac{\gamma}{\lambda} z_{t-1} - \frac{1}{\lambda} \varepsilon_t^{(2)}$$

and combining it with the first one leads to a standard type of difference equation. Substituting then its solution into the transformed third equation from above, one arrives at the bubble-free solution which is given by

$$s_t - s_{t-1} = \left[\frac{\rho - \gamma}{\lambda} \right] z_{t-1} - \frac{1}{\lambda} \varepsilon_t^{(2)} + \left(\frac{1}{\lambda + \gamma - \rho} \right) \varepsilon_t^{(1)} \quad . \quad (1)$$

It follows that model (I) implies an equation in which the parameter before z_{t-1} is "significantly" negative, e.g. if $\gamma > \rho$ and $\lambda \approx 0$. It is this negativity which will be of importance in the next section.

2 History of the Model

In several papers (*H. Garbers*, [3],[4],[5]) the author tries to give a critical discussion of two approaches which study the relationship between forward and spot exchange rates. It starts by testing especially the (efficiency and risk neutrality) restriction $b = 1$ in the framework of the classical regression model

$$s_t = a + b f_{t-1} + \varepsilon_t \quad (2)$$

where s_t is the log of the spot rate in t and f_{t-1} the log of the forward rate in $t - 1$. With the *USD/CHF* relation (monthly data, last working day in Switzerland, 1974.1 – 1988.10) one gets the *following* results from an *OLS* estimation of equation (2):

coefficient	value	standard error
a	0.021	0.010
b	0.970	0.014

Table 1: *Durbin-Watson statistic = 1.88, $R^2 = 0.97$*

Remark 2.1. There is another relationship between s_t and f_t going back to an arbitrage equation between S_t and F_t , the non-logarithmic version of s_t, f_t :

$$S_t(1 + r_t) = (1 + r_t^*)F_t \quad ,$$

where $r_t(r_t^*)$ is the Swiss (US) one month interest rate.

It follows that

$$\frac{F_t}{S_t} = \frac{1 + r_t}{1 + r_t^*}$$

and by taking logarithms using a *Taylor* approximation

$$\begin{aligned} f_t - s_t &\approx r_t - r_t^* (= z_t) \\ f_t &\approx s_t + z_t \quad . \end{aligned}$$

The first and second equation of *B.T. McCallum's* model can therefore be written as (approximately)

$$f_t = E_t[s_{t+1}] + \zeta_t \tag{3}$$

$$\zeta_t = \rho\zeta_{t-1} + \varepsilon_t^{(1)} \quad ; \quad 0 < \rho < 1 \tag{4}$$

decomposing f_t into the sum of two unobserved components with the second one, ζ_t , being $I(0)$. If then f_t contains a unit root, so will $E_t[s_{t+1}]$ and implicitly also s_t . Equation (2) might then, however, be considered as a consistently estimated cointegrated relation with wrong standard errors. We shall come back to this perspective at the end of the paper.

For now, let us consider (2) as representing the conditional distribution of s_t , given f_{t-1} , claiming f_{t-1} to be a weakly exogenous variable with respect to b . Moreover, equation (2) is equivalent to the following equation (5) under the null hypothesis of $b = 1$.

$$s_t - s_{t-1} = a + b(f_{t-1} - s_{t-1}) + \varepsilon_t \quad . \quad (5)$$

Using the same data set as for (2) one obtains the following *OLS* regression results:

coefficient	value	standard error
a	- 0.02	0.006
b	- 3.82	1.113

Table 2: ***Durbin-Watson statistic = 2.04, $R^2 = 0.06$***

According to these results we reject the hypothesis that $b = 1$ in the framework of (5). A possible reaction would be to follow a general suggestion of *E.F. Fama* [2] and to introduce a risk premium into the framework of the model. But this is how *B.T. McCallum's* model (I) comes into the picture. It implies a risk premium and equation (1) follows from it which is close to equation (5), because of $a \approx 0$ and $z_{t-1} \approx f_{t-1} - s_{t-1}$ while the parameter before z_{t-1} can very well be less than -3.8 because of (1).

There is, however, some evidence that *equation (1), like (5), should generally be rejected as misspecified*. By estimating e.g. – in agreement with *E.F. Fama's* [2] maintained hypothesis – the spectral densities $f(\lambda)$ of $s_t - s_{t-1}$ and of $f_t - s_t$ one finds them very different (see figures 1 and 2), while (1) is close to

$$s_t - s_{t-1} = a + bL(f_t - s_t) + \varepsilon_t \quad . \quad (6)$$

But this is a linear time invariant filter, so that

$$\begin{aligned} f_{s_t-s_{t-1}}(\lambda) &= T(\lambda)f_{f_t-s_t}(\lambda) + f_\varepsilon(\lambda) \\ &= b^2 f_{f_t-s_t}(\lambda) + \frac{\sigma_\varepsilon^2}{2\pi} \end{aligned}$$

where the $f(\lambda)$ terms represent spectral densities and

$$T(\lambda) = \left(b \exp(-i\lambda)\right) \left(b \exp(i\lambda)\right) = b^2$$

is the transfer function of the filter.

It follows that equation (1) implies a spectral density for the output which is quite different from the one we "observe". See the following figures. We decide therefore to reject the hypothesis of (1) [and (5)].²

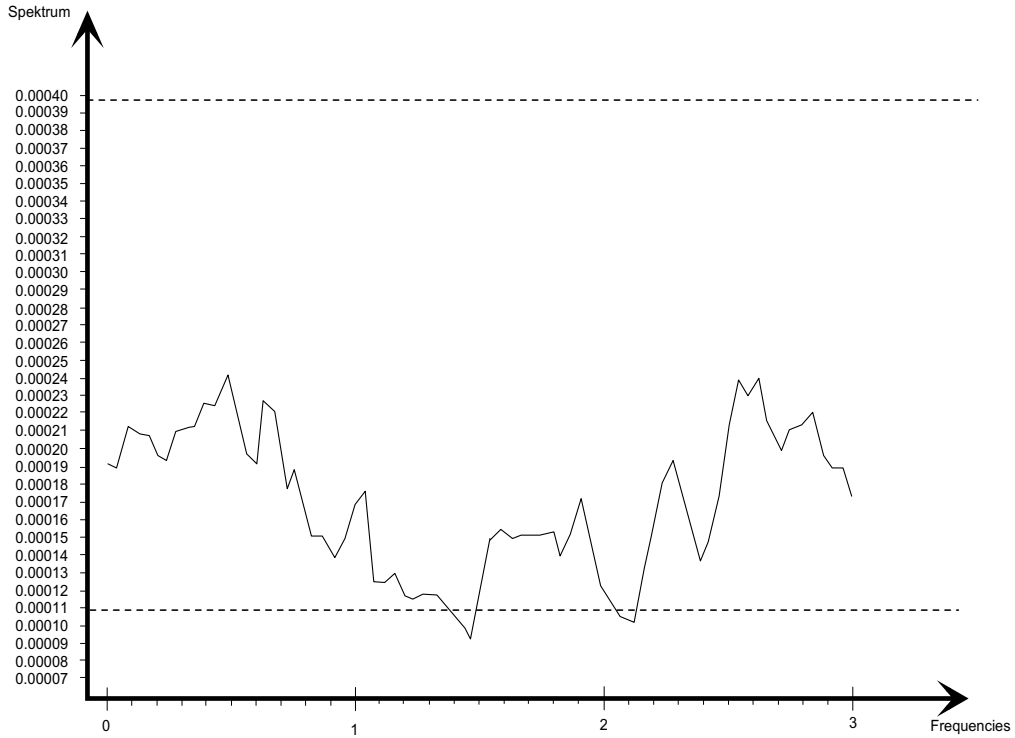


Figure 1: *The estimated spectrum of $s_t - s_{t-1}$, data 1974.1 – 1987.6. The two limiting lines represent the upper and lower limit of a white noise process with the same variance*

² Notice, by the way, that this does not imply the rejection of (2) with $b \neq 1$.

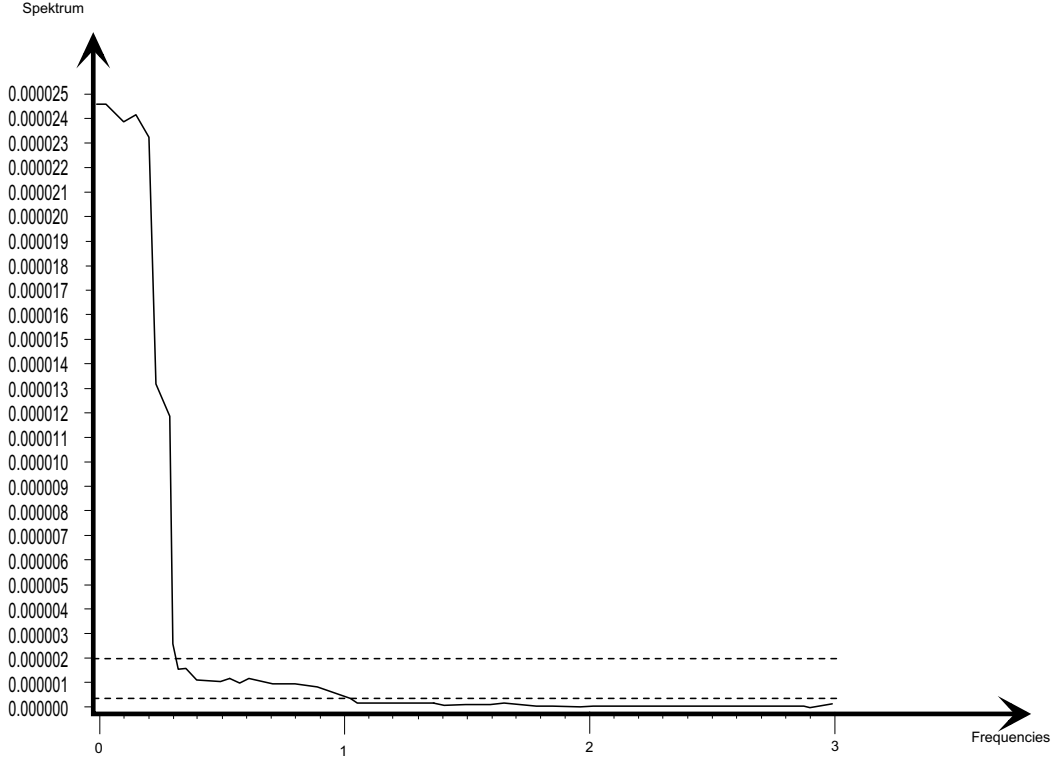


Figure 2: *The estimated spectrum of $f_t - s_t$, data 1974.1 – 1987.6*

3 Rejecting McCallum's Model

As *B.T. McCallum's* model implies (1) as a solution and (1) is an approximation to the rejected filter of (6), the model should be rejected too.

As the model consists of a whole set of hypotheses which are combined by it, is there any single one that can be made responsible for the rejection? If so, we should substitute this one by a "better" hypothesis, so that the modified model does not show the deficient implication any longer.

It is the author's point of view that the conditional expectation $E_t[s_{t+1}]$ which appears in the first equation of (I) does represent a shortcoming of *B.T. McCallum's* model and should be substituted by a different

concept.

As $E_t[s_{t+1}]$ is not directly observed, a new data source has been used in the literature to shed light on the expectation formation of agents: the results of surveys of market participants conducted by financial services firms (see *S. Takagi* [17]). Accordingly, different types of market participants form expectations in different ways, some are more heavily represented at the short horizons (speculators) while others are at the long horizons (investors). The distinction between speculators and investors illustrates that not all participants share the same expectations. We will illustrate that still a lot more different expectations have to be introduced to arrive at a proper model for the *DGP* linking s_{t+1} and f_t .

4 Rational Beliefs, Stable Processes and Endogeneous Uncertainty

M. Kurz [10] was among the first who realized that e.g. rational expectations might be a misleading concept in a non-static world. *He tries to develop an alternative quantitative anchor, starting with Birkhoff's theorem.* To introduce and to generalize it some preparatory remarks have to be made. They will allow us afterwards to introduce *M. Kurz's* concept of a rational belief.

Given a probability space (Ω, \mathcal{A}, P) and a measure preserving transformation $T : \Omega \rightarrow \Omega$. If for certain $A \in \mathcal{A}$, it is true that $T^{-1}(A) = A$, then A is referred to as invariant under T . The system \mathcal{I} of all invariant sets of \mathcal{A} constitutes a σ algebra. A measure preserving transformation T is referred to as ergodic, if for each invariant set A , it is true that either $P(A) = 1$ or $P(A) = 0$. After these preparations we can now formulate the theorem mentioned above.

Theorem (Birkhoff, see *A.N. Shiryaev*, p. 381 [15])*

Assuming T is a measure preserving transformation and X is a random variable defined on (Ω, \mathcal{A}, P) with $E|X| < \infty$. With probability 1 it is then true that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} X(T^j(\omega)) = E[X|\mathcal{I}] \quad .$$

*We omit a second statement of the theorem.

M. Kurz then tried to "invert" the theorem by starting with the statement of the theorem asking for the class of stochastic processes, which imply it.

Assuming $\vec{x}_t \in \mathbb{R}^n$ is a vector of n observed variables and $\vec{x} = (\vec{x}_0, \vec{x}_1, \dots)$ is an infinite sequence in

$$(\mathbb{R}^n)^\infty := \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \times \dots \quad .$$

He then sets $\Omega = (\mathbb{R}^n)^\infty$ and take as a σ algebra \mathcal{B} the one generated from the *Borel* sets. In addition $\{X_t, t = 0, 1, 2, \dots\}$ is assumed to be a stochastic process, Q the probability measure induced by the process, but not known and (Ω, \mathcal{B}, Q) the true probability space. Each market participant $k \in \{1, \dots, K\}$ attempts then, according to *M. Kurz*, to "track down" this true probability space through the observations $(\vec{x}_0, \vec{x}_1, \dots, \vec{x}_t)$. In this way he constructs a probability space $(\Omega, \mathcal{B}, Q_k)$, whereby Q_k represents his beliefs concerning Q , from which it will usually deviate.

Assuming T (hereafter) the so-called shift operator, which is defined by

$$T(\vec{x}_t, \vec{x}_{t+1}, \vec{x}_{t+2}, \dots) = (\vec{x}_{t+1}, \vec{x}_{t+2}, \dots) \quad ,$$

this means, amongst other things, that T and all $T^j, j = 1, 2, \dots$ are functions from $\Omega \rightarrow \Omega$. T is then generally not a measure preserving transformation. If $B \in \mathcal{B}$ is then a finite dimensional cylinder set, then

$$m_n(B)(x) := \frac{1}{n} \sum_{l=0}^{n-1} 1_B(T^l(x))$$

with

$$1_B(y) := \begin{cases} 1 & ; \quad \text{if } y \in B \\ 0 & ; \quad \text{if } y \notin B \end{cases}$$

and

$$x := (x_0, x_1, x_2, \dots)$$

is the relative frequency of the occurrence of B since $t = 0$.

Each of the K individuals can in principle calculate $m_n(B)(x)$ from the given time series.

A dynamic system $(\Omega, \mathcal{B}, Q, T)$ is referred to as *stable*, if for each $B \in \mathcal{B}$ of finite dimensions, it can be stated that:

$$\lim_{n \rightarrow \infty} m_n(B)(x) =: m(B)(x) \quad \text{with } Q = 1 \quad .$$

It applies that under certain conditions $m(B)(x)$, which we will take as given, can be extended to a uniquely defined probability measure m on (Ω, \mathcal{B}) , which is independent of x . In this way $(\Omega, \mathcal{B}, m, T)$ constitutes a (strictly) stationary process, where possibly $m \neq Q$ and the agents can only learn the stationary component m of Q .

On the other hand, $(\Omega, \mathcal{B}, Q, T)$ is referred to as **weakly asymptotically mean stationary (WAMS)**, if for each of these $B \in \mathcal{B}$, it

can be stated that

$$m^Q(B) := \lim \frac{1}{n} \sum_{l=0}^{n-1} Q(T^{-l}(B))$$

is well-defined. It should be noted that $m^Q(B)$ represents a theoretical expression, deduced alone from the dynamic system. It shows no relationship to the data x . Its analytical value becomes apparent through the following

Theorem (Theorem 1 in *M. Kurz*, p. 15 [10]):

$(\Omega, \mathcal{B}, Q, T)$ is stable if and only if it is a *WAMS*. It is also true that

$$m(B) = m^Q(B) \quad ; \quad \forall B \in \mathcal{B} \quad .$$

Overall it can then be deduced from the previous arguments, that each stable dynamic system generates a uniquely determined stationary probability measure $m^Q(\cdot)$, *which can in principle be learned by all agents from the data*.

A belief Q_k of an agent is finally referred to as *rational relative to m* , if

1. it cannot be contradicted by the data that is the dynamic system $(\Omega, \mathcal{B}, Q_k, T)$ is stable and $m^{Q_k} = m$.
2. from $m(B) > 0$ it always follows that $Q_k(B) > 0$.

As a rule, a rational belief is characterised by more than just its stationary implications. A further component of Q_k is a measure Q_k^\perp , which is orthogonal to $m(\cdot)$, i.e. there is a $B \in \mathcal{B}$ and a $B^c := \Omega - B$ so that

$$m(B) = 1, m(B^c) = 0 \quad \text{and} \quad Q_k^\perp(B) = 0, Q_k^\perp(B^c) = 1 \quad .$$

Indeed according to *M. Kurz*, the following theorem is valid:

Theorem (Theorem 2 in *M. Kurz*, p. 16 [10])*

If a *DGP* behaves according to a stable dynamic system $(\Omega, \mathcal{B}, Q, T)$, then it implies for each rational belief Q_k that

1. $Q_k = \lambda_k Q_a + (1 - \lambda_k) Q_k^\perp$ with $0 < \lambda_k \leq 1$, whereby
2. Q_a is equivalent to m and Q_k^\perp is orthogonal to m .

It also implies that both $(\Omega, \mathcal{B}, Q_a, T)$ and $(\Omega, \mathcal{B}, Q_k^\perp, T)$ are stable with $m^{Q_a} = m^{Q_k^\perp} = m$.

*We ignore a second statement of the theorem.

Notice that we have obtained a probability distribution which represents a *subjective quantity*. The subjective components of the distribution being λ_k and Q_k^\perp . They are linked to the system's non-stationarity and will generally induce heterogeneous (but "rational") beliefs amongst the agents.

Realizing then that there is a whole set $B(Q)$ of rational beliefs relative to m , agents make, according to *M. Kurz*, at some time their choices out of $B(Q)$. Allowing then these choices to be cross-correlated *M. Kurz* demonstrates that even simple *Monte Carlo* studies generate time series showing a surprisingly "realistic" behavior. *M. Kurz* [10] [11] refers to an *endogenous uncertainty*: Agents select their subjective probabilities from $B(Q)$ changing their choices from time to time according to fairly unknown principles of some social practices in communities (see *H. Garbers* [6]). Endogeneous uncertainty is then expected to generate switching regimes in the corresponding variables.

5 Switching Regimes in s_t

We proceed by presenting some evidence that switching regimes (and autoregressive structures) can indeed be found in s_t . Just for this reason we apply provisionally a *segmented random walk* model to s_t according to which at each time period t , s_t is in state x_t , being 1 or 0, where x_t is not directly observable. With a time invariant probability of p_{11} it will be in state 1 in period t if it was in state 1 in period $t-1$, while it will change to state 0 in t with a probability of $1-p_{11}$. For short we write:

$$\begin{aligned} P(x_t = 1 \mid x_{t-1} = 0) &= 1 - p_{11} \\ P(x_t = 0 \mid x_{t-1} = 0) &= p_{11} \\ P(x_t = 1 \mid x_{t-1} = 1) &= p_{00} \\ P(x_t = 0 \mid x_{t-1} = 1) &= 1 - p_{00} \quad . \end{aligned}$$

The segmented random walk model can then be represented by

$$s_t - s_{t-1} = \mu_1 + (\mu_2 - \mu_1) x_t + \varepsilon_t^{(1)} + (\varepsilon_t^{(2)} - \varepsilon_t^{(1)}) x_t \quad , \quad (7)$$

where $x_t \in \{0, 1\}$ and $\varepsilon_t^{(1)}, \varepsilon_t^{(2)}$ are independent normal white noise processes. The probability distribution of x_t behaves according to a *Markoff* chain:

$$\begin{pmatrix} q_t & (1 - q_t) \end{pmatrix} = \begin{pmatrix} q_{t-1} & (1 - q_{t-1}) \end{pmatrix} \begin{pmatrix} p_{11} & (1 - p_{11}) \\ (1 - p_{00}) & p_{00} \end{pmatrix} \quad , \quad (8)$$

where

$$q_t \equiv P[x_t = 0] \quad .$$

It follows for $1 > p_{11} > 0$ and $1 > p_{00} > 0$ that the distribution of

x_t converges to a well known limit distribution so that

$$x_\infty \equiv \begin{cases} 0 & \text{with probability } q \\ 1 & \text{with probability } 1 - q \end{cases} ,$$

where

$$q = \frac{1 - p_{00}}{2 - p_{11} - p_{00}} .$$

p_{11}, p_{00} and the two drift terms μ_1, μ_2 are unknown and have to be estimated. Moreover, the two states $x_t = 0, x_t = 1$ may not only differ in respect to the drift term μ but also in respect to the variance σ^2 which is also unknown. There are, therefore, six population parameters,

$$\vec{\Theta} := \left(\mu_1 , \mu_2 , \sigma_1 , \sigma_2 , p_{11} , p_{00} \right)'$$

which determine (with a corresponding density function) the distribution of s_t , given x_t , and the distribution of x_t , given x_{t-1} .

Taking as a first example end of the month data for the period of 1983.1 - 1996.6 we arrive at the following point estimates $\hat{\vec{\Theta}}$ for $\vec{\Theta}$ together with the corresponding standard errors (in brackets)³:

$\hat{\mu}_1 [\cdot 10^2]$	$\hat{\mu}_2 [\cdot 10^2]$	$\hat{\sigma}_1^2 [\cdot 10^4]$	$\hat{\sigma}_2^2 [\cdot 10^4]$	\hat{p}_{11}	\hat{p}_{00}
-0.54	1.11	14.49	4.92	0.99	0.98
(0.34)	(0.49)	(1.77)	(1.62)	(0.01)	(0.03)

Table 3: ***Estimation results of the segmented random walk s_t***

The two states of s_t differ remarkably as to μ and σ^2 , the drift term is important and differs in sign between the states while the variances imply large deviation between the risk premia of state 1 and 0.

³ An extended data set (1980.2 - 2003.6, for example) will be analyzed in Part II of the paper allowing up to 3 different regimes. The results appear to be numerically sensitive to alterations in the time frame like those from table 3. But there are no doubts about the existence of different regimes.

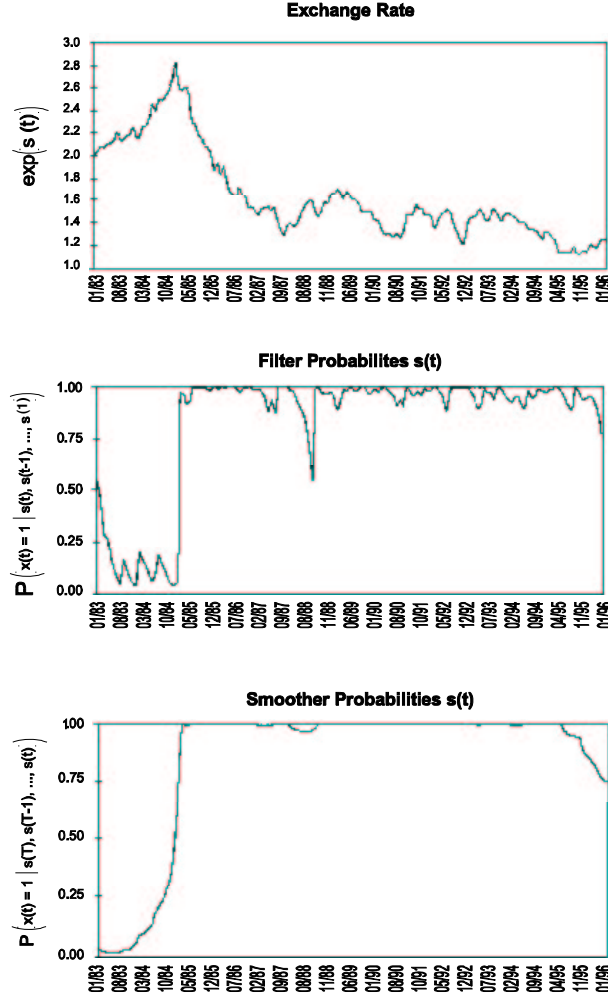


Figure 3: *Original time series and estimated probabilities for the segmented random walk of s_t*

Moreover, we include in our results the graphs of figure 3. The latter two represent *the filter and the smoother probabilities of state 1* (see *J.D. Hamilton*, chapter 22 [7]). According to these figures s_t was most of the time in state 1 of a declining *USD* with a high variance. s_t was in state 0 during 1983 until the beginning of 1985.

Notice, finally, that the point estimates \hat{p}_{11} of p_{11} and \hat{p}_{00} of p_{00} are 0.99 and 0.98 respectively with a standard error of 0.01, 0.03. Given that and the evidence for the existence of two regimes, we accept the hypothesis that the *Markov* chain is non-ergodic and that s_t is not a *WAMS* and therefore not a stable process.

6 A few conclusions

- a) Applying his theory to given sets of general time series, *M. Kurz* [10] is supplementing it by telling us that:

”The data may need to be ”cleaned” for trend and deterministic cycles if ”stability” is to be a useful tool. This is standard practice in time series analysis (...).”

However, this underestimates the importance of a major problem area. To demonstrate, remember that in time series analysis one comes across data structures since the famous work of *C.R. Nelson and C.I. Plosser* [13], according to which macro-economic variables X_t exhibit predominantly stochastic trends and additionally moving averages⁴:

$$(1 - L)X_t = \mu + B(L)\varepsilon(t) \quad .$$

Using a recursive solution algorithm for this stochastic difference equation, taking $X_0 = 0$ and $\varepsilon_t = 0$ for $t \leq 0$, leads to a representation of X_t according to

$$\begin{aligned} X_t &= \tau_t + C[L] \varepsilon_t \\ \tau_t &= \tau_{t-1} + \mu + C[1] \varepsilon_t \quad . \end{aligned}$$

with some lag polynomial $C[L]$. It follows that the innovation of the random walk component is *perfectly correlated* with the innovation of the stationary component of X_t .

⁴ *J.H. Stock and M.W. Watson* [16] later investigated the multivariate implications of these univariate results and thereby arrived at their famous ”common trend” model:

$$\begin{aligned} \vec{X}_t &= \vec{\gamma} + A\vec{\tau}_t + D(L)\vec{\varepsilon}_t \\ \vec{\tau}_t &= \vec{\mu} + \vec{\tau}_{t-1} + \eta_t \end{aligned}$$

Allowing for independence of the innovations one arrives at a first model in state space form:

$$\begin{aligned} X_t &= \omega_t + \eta_t \\ \omega_t &= \omega_{t-1} + a_0 + \delta_t \quad , \end{aligned}$$

where $(\eta_t)_{t \in \mathbb{Z}}$ and $(\delta_t)_{t \in \mathbb{Z}}$ are *independent* white noise processes, $a_0 \in \mathbb{R}$, a_0 fix.

Allowing additionally for a certain time variability of a_0 one gets

$$\begin{aligned} X_t &= \omega_t + \eta_t \\ \omega_t &= \omega_{t-1} + a_t + \delta_t \\ a_t &= a_{t-1} + v_t \quad , \end{aligned}$$

where $(\eta_t)_{t \in \mathbb{Z}}$, $(\delta_t)_{t \in \mathbb{Z}}$ and $(v_t)_{t \in \mathbb{Z}}$ are independent "white noise" processes. It follows, again by a recursive solution algorithm, that

$$X_t = (\eta_t - \eta_0) + \sum_{i=1}^t \delta_i + t(a_0 + v_1) + (t-1)v_2 + \dots + v_t$$

and *the parameters* with respect to $t, t-1, \dots, 1$ are *random variables*. Obviously, X_t has a complex stochastic structure and there seems to be no "standard practice in time series analysis" to eliminate e.g. the trend "in t ".

- b) Applying even to a *Nelson-Plosser* type of *DGP* a "standard practice in time series analysis", for example, a *Hodrick-Prescott* filter, there is a considerable *danger of generating pure artefacts* (*A. Jaeger* [9]). Notice, a filter is an operator which transforms stochastic processes into stochastic processes. And there is a well developed theory for the case of linear filters, which are defined on spaces of weakly stationary stochastic processes with absolutely

summable covariance functions. However, the *Hodrick-Prescott* filter is a typical procedure used to filter time sequences of a *DGP* comprising various components which are predominantly located in *different frequency bands*, while the *DGP of exchange rates* e.g. consists of various regimes which are manifested in *overlapping frequency bands*.

- c) Using a cointegration approach for s_t (and f_t) "as a standard practice" instead of a prefiltering procedure does again lead to problems:

Taking monthly data for s_t and f_t , from 1975.1 until 2003.5, and using a **Phillips-Perron** t type test⁵ the null hypothesis "There is a unit root" is accepted with a

$$p \text{ value of } 0.2881 \text{ (for } s_t) \quad \text{and} \quad 0.2961 \text{ (for } f_t) \quad .$$

Using an **Augmented Dickey-Fuller** test with different lags one constantly receives the same results.

Testing the null hypothesis "stationary around a constant" against the alternative of a unit root according to a *KPSS* test⁶, the null is rejected at the 1 % level for s_t as well as for f_t .

We proceed by asking whether s_t and f_t are cointegrated so that $s_t - f_t$ is an $I(0)$ variable. Considering the null hypothesis that this difference contains a unit root, we get a p value of

⁵ with a *Bartlett* window of bandwidth 5

⁶ **Kwiatkowski, Phillips, Schmidt and Shin** test, see Zivot, E., Wang, J.: *Modeling Financial Time Series with S-Plus*; Insightful Corporation, 2003, p. 123ff

0.1619 by a PP test
0.1204 by an ADF test (1 lag)
0.0468 (2 lags)
0.0816 (3 lags) .

The conclusion is that s_t and f_t are not cointegrated with a cointegrating vector $(1, -1)$.

Instead of taking the a priori given cointegrating vector we postulate an unknown normalized vector in order to estimate it and test its existence in a next step using the *PP* test and the *ADF* test as adapted by *Phillips and Ouliaris*.

Analyzing, then, the *OLS* residuals of a linear regression of s_t on f_t (including a constant term) we *accept the unit root hypothesis for the residuals* using a *PP* test (p value of 0.138) and an *ADF* test (taking one, or more, lags) with a p value of 0.1511 (1 lag), 0.0915 (2 lags) and even higher p values in case of more lags.

The conclusion is that s_t and f_t are not cointegrated at all, although both time series are very similar: The *OLS* estimate of the f_t parameter is 0.9950 with an estimated standard error of 0.007.

Remark 6.1. The *OLS* estimator of an (eventually existing) cointegration vector and especially the standard errors could be improved by using *J.H. Stock and M.W. Watson's* **Dynamic *OLS*** estimator (see *E. Zivot and J. Wang* [18], page 435). Applying the *DOLS* procedure with a lag and a lead of 3 to the same data set as before, the f_t parameter is estimated at 0.9945 with a standard error of 0.0013.

- d) It does not come as a surprise that s_t and f_t are not cointegrated, as this would have implied the existence of a common trend τ_t in the framework of $I(1)$ processes:

$$\begin{aligned} s_t &= \tau_t + u_t \\ f_t &= \tau_t + v_t \\ \tau_t &= \tau_{t-1} + w_t \quad , \end{aligned}$$

where u_t and v_t are weakly stationary processes and w_t is white noise. But taking the *segmented* random walk character of s_t as given, one would at least need something like *D.F. Hendry's* theory of co-breaking [8] for the analysis of the system (s_t, f_t) .

- e) However, we consider the non-linearities in the relations between s_t and additional variables as a *basic problem*. Note that the rational belief concept does not appear to be operational. It addresses however a few important points like the subjectivity of the probabilities implied, the endogenous uncertainties, and the switching regimes giving way to the discussion of processes and practices in the background.

Remember then, *B.T. McCallum's* model (*I*) shows some additional structure which we ignored up to now. It contains a feedback relation and an unobserved variable, (ζ_t) , which represents, amongst others, a set of fundamental variables discussed in exchange rate theory. It includes for instance the relative growth rate of the gross national product, the difference in inflation rates, the difference in the short-term and long-term return, the productivity ratings of the two countries, etc. Economic theories emphasize, however, that at any time it is *only a non-anticipated part of the variables* – an innovation – which affects the exchange rate. The following description of a foreign exchange market by *P. De Grauwe* [1] refers to these theories:

He observed that since its introduction up to May 2000, the *Euro* has lost 25 % of its value in comparison with the *USD*. He analyzed the drop in exchange rate and diagnosed that it had practically no relationship with new information which has become available concerning the basic fundamental variables. *P. De Grauwe* then states

”(...) that it is not the news in the fundamentals that drives the exchange rate changes, but rather the other way around: changes in the exchange rate lead to a selection of news about the fundamentals (present and future) that is consistent with the observed exchange rate changes. (...)

There is great uncertainty among economists about how fundamentals affect the exchange rate. (...). Because we are so uncertain about the underlying fundamentals and their impact, the exchange rate movements themselves become a signal to search for those fundamental variables that can explain the particular exchange rate movement. Thus, when at the start of 1999 the dollar started to move upwards, this became a signal of fundamental strength of the US economy and fundamental weakness of Euroland's economy. This set in motion a search for good news about America and bad news about Europe (...) creating (positive) beliefs about the US (...) and negative beliefs about the European economy ... (which) reinforced the exchange rate movements (...). Obviously this process of creating (...) beliefs can only go on as long as the facts are not too inconsistent with (them).”

The question then is: Why should this story be of special importance for modeling foreign exchange markets? Clearly, if it is important, *complex cognitive processes and social practices come*

into the picture. But what about agents' rationality in it? In order to substitute $E_t[s_{t+1}]$ and to proceed with our misspecification analysis of *Mc Callum's* model, further research is needed and will be presented in Part II of this paper.

References

- [1] **De Grauwe, Paul:** *Exchange Rates in Search of Fundamentals: The Case of the Euro-Dollar Rate*; International Finance, Volume 3, Issue 3, 2000, pp. 329 - 356
- [2] **Fama, Eugene F.:** *Forward and Spot Exchange Rates*; In: Journal of Monetary Economics 14, 1984, S. 319-338
- [3] **Garbers, H.:** *A Misspecification Analysis of the Relationship between Spot and Forward Exchange Rates*; European Economic Review, 1987, 31, pp. 1407-1417
- [4] **Garbers, H.:** *The relationship between Exchange Rates and Interest Rates: The Swiss Franc Case*; Wirtschaftswissenschaftliche Beiträge, vol. 60, 1992, Physica Heidelberg
- [5] **Garbers, H.:** *Constructing an Empirical Model for Swiss Franc Exchange Rates and Interest Rate Differentials*; Studies in Empirical Economics, 1994, Physica Heidelberg
- [6] **Garbers, H.:** *The Rationality of Economic Agents: Social Practices in Foreign Exchange and Stock Markets*; unpublished paper, Zurich 2002
- [7] **Hamilton, J.D.:** *Time Series Analysis*; 1994, Princeton University Press
- [8] **Hendry, D.F.:** *A Theory of Co-breaking*; 1996, Working paper, University of Oxford
- [9] **Jaeger, A.:** *Mechanical Detrending by Hodrick-Prescott Filtering: A Note*; In: Empirical Economics (1994) 19: pp. 493-500
- [10] **Kurz, M.:** *Endogeneous Economic Fluctuations*; Springer Heidelberg, New York, 1997
- [11] **Kurz, M. (Ed.):** *Endogenous Uncertainty: A unified View of Market Volatility*, Manuscript; Stanford University, 1998
- [12] **McCallum, B.T.:** *A Reconsideration of the uncovered Interest Parity Relationship*; Journal of Monetary Economics, 1994, 33, pp. 105-132
- [13] **Nelson, C.R., Plosser, C.I.:** *Trends and Random Walks in Macroeconomic Time Series*; Journal of Monetary Economics, 1982, pp. 139-162
- [14] **Shiller, Robert J.:** *Irrational Exuberance*; Princeton University Press, 2000
- [15] **Shiryayew, A.N.:** *Probability*; New York, Berlin, Heidelberg, Tokyo, 1984
- [16] **Stock, J.H., Watson, M.W.:** *Testing for Common Trends*; Journal of the American Statistical Association, 1988, pp. 1097-1107
- [17] **Takagi, S.:** *Exchange Rate Expectations: A Survey of Survey Studies*; International Monetary Fund Staff Paper 38, 1991, pp. 156-183
- [18] **Zivot, E., Wang, J.:** *Modeling Financial Time Series with S-Plus*; Insightful Corporation, 2003